Articles

Molecular Characterization of Two Galactosemia Mutations and One Polymorphism: Implications for Structure–Function Analysis of Human Galactose-1-phosphate Uridyltransferase

Juergen K. V. Reichardt,*,‡ Harvey L. Levy,§ and Savio L. C. Woo‡

Howard Hughes Medical Institute and Department of Cell Biology, Baylor College of Medicine, Texas Medical Center, Houston, Texas 77030-3498, and PKU Program, Children's Hospital, Boston, Massachusetts 02115

Received November 20, 1991; Revised Manuscript Received March 11, 1992

ABSTRACT: We report here the molecular characterization of two galactosemia mutations, L74P and F171S, and one polymorphism, S135L, in human galactose-1-phosphate uridyltransferase (GALT). Both galactosemia mutations result in reduced enzymatic activity when reconstructed in the cDNA and overexpressed. The polymorphism, in contrast, has near normal activity. Both mutations affect evolutionarily conserved residues, suggesting that they are functionally important, while the polymorphism occurs in a nonconserved domain which is presumably not critical for enzymatic function. The F171S mutation is close to the putative active-site nucleophile. Our data further support the notion of molecular heterogeneity of galactosemia and suggest that galactosemia mutations and GALT polymorphisms may be useful tools in highlighting different functional domains in human GALT.

Classic galactosemia (McKusick 230400) results from deficiency of the ubiquitously expressed enzyme galactose-1-phosphate uridyltransferase (GALT;[1] EC 2.7.7.12; Segal, 1989). Afflicted infants suffer from vomiting and diarrhea after milk ingestion. If untreated, they will develop cataracts, hepatosplenomegaly, failure to thrive, and *Escherichia coli* sepsis. These symptoms can be avoided by placing patients as soon as possible on a galactose-restricted diet (Segal, 1989). Therefore, many U.S. states and Western countries have instituted newborn screening programs. Unfortunately, most well-treated patients will suffer from neurologic abnormalities and ovarian failure (Waggoner et al., 1990). Galactosemia is an autosomal recessive disorder with a frequency of about 1:60000 newborns (Levy & Hammersen, 1978).

GALT catalyzes the conversion of galactose 1-phosphate and uridine diphosphoglucose (UDPG) to uridine diphosphogalactose (UDP-gal) and glucose 1-phosphate. Its absence in galactosemia results in elevated levels of galactose, galactose 1-phosphate, and other metabolites in the untreated state (Segal, 1989). It is thought that these compounds result in the early-onset complications which can be avoided by a galactose-restricted diet. The late-onset symptoms may result from a UDP-gal deficit first described by Kaufman et al. (1988). This deficit is the direct result of a deficiency of GALT and is reflected in subnormal galactosylation of glycoproteins and glycolipids (Dobie et al., 1990; Petry et al., 1991).

Through the work of Frey and co-workers, much has been learned about the *E. coli* enzyme. This enzyme, galT, undergoes a ping-pong reaction mechanism with a covalent uridylate reaction intermediate derived from UDPG (Frey et al., 1982; Field et al., 1988). Furthermore, they were able to show that the covalent uridylate intermediate is bound to *E. coli* histidine-166 (Kim et al., 1990). It was proposed on the basis of sequence similarities between the human, yeast, and *E. coli* enzymes that they all use similar intermediates (Reichardt & Berg, 1988b).

The molecular analysis of galactosemia began with the cloning and characterization of a full-length, expressible GALT cDNA (Reichardt & Berg, 1988a; Flach et al., 1990). Studies of the GALT gene in galactosemic patients based on Southern (Reichardt, 1989, 1991), Northern (Reichardt, 1989, 1991), and Western blotting experiments (Banroques et al., 1983; Reichardt, 1989, 1991) led to the suggestion that galactosemia is caused predominantly by missense mutations (Reichardt, 1989, 1991). In fact, six disease-causing missense mutations have been characterized to date (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992). Five of the mutations occurred in evolutionarily conserved domains of GALT, suggesting that the affected residues are of functional importance. Both published polymorphisms, in contrast, affected variable residues (Reichardt & Woo, 1991). Therefore, it may be possible to use the mutational analysis of human galactosemia to analyze functional domains in human GALT. The large number of different mutations causing galactosemia suggested furthermore that this disease is heterogeneous at the molecular level (Reichardt & Woo, 1991; Reichardt et al., 1991).

Here we report two galactosemia missense mutations, L74P and F171S, that highlight functionally important domains of the enzyme. The polymorphism, S135L, was found at a variable position in GALT. It is noteworthy that the F171S mutation occurs near the putative active-site nucleophile. Both mutations result in severe losses of GALT specific activity

* Address correspondence to this author.
‡ Baylor College of Medicine.
§ Children's Hospital.

[1] Abbreviations: ASO, allele-specific oligonucleotide; GALT, galactose-1-phosphate uridyltransferase; PCR, polymerase chain reaction; PKU, phenylketonuria; UDPG, uridine diphosphoglucose; UDP-gal, uridine diphosphogalactose.

0006-2960/92/0431-5430$03.00/0

while the polymorphism encodes almost normal activity. Finally, all three substitutions are rare among our patients. These data further support the notion of molecular heterogeneity in galactosemia. Our data also suggest that galactosemia mutations will be a useful guide in a future rigorous structure–function analysis of human GALT by highlighting important domains. Polymorphisms, in contrast, mark less critical residues.

Materials and Methods

Cell Culture. Lymphoblastoid lines were established from two galactosemic patients by Epstein–Barr virus (EBV) transformation of lymphocytes obtained from whole blood (Reichardt & Woo, 1991). Cells were grown in RPMI1640 (Hazleton, Lenexa, KS) supplemented with 10% bovine serum (HyClone, Logan, UT). Two cell lines were obtained, JR140 and 154. Patients were screened and diagnosed by a positive newborn metabolite test and by assaying their red cell GALT activity. They were identified through the New England Regional Newborn Screening Program in Boston and have been continuously followed on treatment. The GM2796 line was obtained from the NIGMS Human Mutant Cell Repository (Camden, NJ). *cos* cells were grown in HGDMEM supplemented with 10% bovine serum.

Amplification and Sequencing. Cellular RNA was prepared from log-phase cells by centrifugation through a CsCl cushion (Chirgwin et al., 1979). Total cDNA libraries were synthesized from poly(A^+) RNA and amplified exactly as described (Reichardt & Woo, 1991). The entire coding region of GALT was amplified with 2.5 units of *Taq* DNA polymerase (Promega, Madison, WI) exactly as detailed elsewhere (Reichardt & Woo, 1991). After gel purification to remove the PCR primers, amplified products were directly sequenced with Sequenase 2.0 (USB, Cleveland, OH) using internal primers (Reichardt & Woo, 1991; Reichardt et al., 1991).

ASO (Allele-Specific Oligonucleotide) Screening and Expression Analysis. 19mers spanning each mutation (ASI, Pleasanton, CA; L74P, TCAACCCTCt/cGTGTCCTGG; S135L, ACCCCTGGTc/tGGATGTAAC; F171S, TGCAGATCTt/cTGAAAACAA; upper case letters denote identical bases, and lower case letters indicate mutant bases) were labeled with [γ-^{32}P]ATP (DuPont, Boston, MA) by T4 polynucleotide kinase (NEB, Beverly, MA). These probes were then hybridized to amplified cDNAs from 14 normal and 26 patient individuals (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992) overnight in 6× SSPE, 0.5% SDS, 1× Denhardt's, and 200 μg of denatured salmon sperm DNA at 42 °C. After four washes at room temperature, a stringency wash was performed at 63 °C in 3 M tetramethylammonium chloride (Aldrich, Milwaukee, WI; Wood et al., 1985; Reichardt & Woo, 1991), 0.5% SDS, 10 mM Tris, and 1 mM EDTA, pH 7.5, for 10 min.

Nucleotide substitutions were reconstructed for expression analysis by site-directed mutagenesis with the same 19mers in uracil-containing pJR16 (Reichardt & Woo, 1991). Constructs were confirmed by sequence analysis in the relevant region from rescued single-stranded DNA. Phagemid DNAs were electroporated into *cos* cells and assayed for enzyme activity and immunoreactive protein as described previously 72-h posttransfection (Reichardt & Woo, 1991).

Computer Analysis. Mutant cDNA sequences were analyzed by the MAP algorithm as implemented by the Molecular Biology Information Resource (MBIR; Baylor College of Medicine). The default 252 restriction recognition sites were loaded.

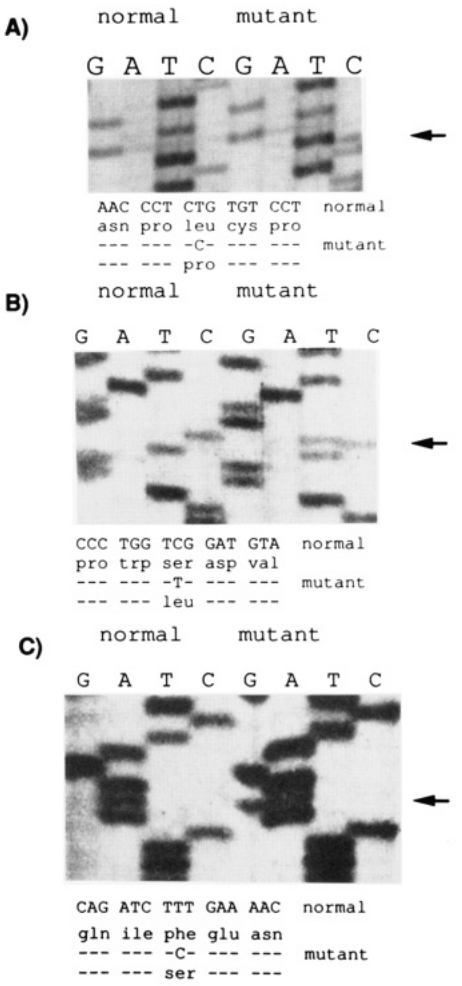

FIGURE 1: Identification of three GALT mutations. Amplified cDNA was sequenced as described under Materials and Methods. Arrows denote the position of each substitution. Panel A is the L74P substitution (found in cell line JR140), panel B shows the S135L substitution, and panel C is the F171S substitution (both of these two substitutions were found in cell line JR154). The sequence in panel C was obtained from the antisense strand.

Results

Mutation Identification. We sequenced amplified cDNA derived from two galactosemic cell lines, JR140 and 154, directly to control for possible PCR (polymerase chain reaction) artifacts and found them to be compound heterozygotes for the common Q188R mutation (data not shown; Reichardt et al., 1991). Three additional substitutions were identified: (1) a T to C transition of nucleotide 249 resulting in the replacement of leucine-74 with proline in JR 140 (L74P; Figure 1A); (2) a C to T transition at base pair 432 substituting serine-135 with leucine (S135L; Figure 1B); (3) a T to C transition of nucleotide 540 replacing phenylalanine-171 with serine (F171S; Figure 1C). These last two substitutions were found in line JR154.

Expression Analysis. We reconstructed each substitution by site-directed mutagenesis in the expression phagemid pJR16 which allows for high-level expression of human GALT in *cos* monkey cells (Reichardt & Woo, 1991). Introduction of the normal and mutant constructs resulted in a 15–31-fold overexpression of immunoreactive protein of the expected size (Table I and data not shown). It is noteworthy that all three substitutions produced near normal amounts of immunoreactive protein when compared to the normal construct, pJR16 (Table I). However, the mutant polypeptides encoded different specific activities (Table I). The L74P mutation (in

Table I: Biochemical Characterization of Two Galactosemia Mutations and One GALT Polymorphism[a]

plasmid	act. (μmol/h)	immunoreact. (^{125}I cpm)	sp act. (μmol h^{-1} cpm^{-1})
	Experiment 1		
none (mock)	23	49	0.47
pJR16 (normal)	729	1325	0.55
pJR26 (135L)	786	1539	0.51
pJR27 (171S)	68	1249	0.04
	Experiment 2		
none (mock)	37	58	0.63
pJR16 (normal)	640	926	0.69
pJR31 (74P)	29	874	0

[a] Transfections with the pJR constructs were corrected for background (mock) activity. Experiments 1 and 2 were performed with different batches of ^{125}I-protein A with slightly different specific radioactivities, resulting in different specific enzyme activities. Enzyme activity and immunoreactivity were determined as published (Reichardt & Woo, 1991).

Table II: Frequency of Two Galactosemia Mutations and One GALT Polymorphism[a]

substitution	normal alleles (positives/total)	galactosemic alleles (positives/total)
L74P	0/14	1/26
S135L	0/14	2/26[b]
F171S	0/14	2/26[b]

[a] The samples for this study are described in this paper and three previous publications (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992). [b] Both of these GALT mutations were found in the same two samples, cell lines JR154 and GM2796, suggesting that they are linked.

pJR31) was found to have no detectable GALT activity while the F171S mutation (in pJR27) resulted in about 7% of the normal specific activity when overexpressed (Table I). In contrast, the S135L substitution (in pJR26) produced near normal levels of enzymatic activity.

Mutation Frequency. We screened our panel of 14 normal and 26 galactosemic cDNAs by allele-specific oligonucleotide (ASO) hybridization to determine the frequency of each substitution (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992). All three GALT mutations are rare: the L74P mutation was found only in the patient we sequenced while the S135L polymorphism and F171S were both also found in a second patient cell line, GM2796, which was obtained from an African–American patient (Table II).

Restriction Sites Altered by GALT Mutations. We analyzed all three substitutions for possible alterations of restriction sites by the MAP algorithm. None of the substitutions reported here are predicted to alter any common sites in the cDNA.

Homology Analysis. We compared the sequences surrounding all three substitutions in the homologous enzymes galT from *E. coli* (Lemaire & Mueller-Hill, 1986), GAL7 from bakers' yeast (Tajima et al., 1985), and GALT from humans (Flach et al., 1990). These three organisms are representatives of such distantly related groups as prokaryotes, unicellular eukaryotes, and multicellular eukaryotes, and the overall sequence identity for the three enzymes is only 35%. Leucine-74 is conserved in all three species and is flanked by an equally well-conserved cysteine, and three amino acids downstream, a conserved glycine is found (Figure 2). Phenylalanine-171 is in an even more impressively conserved domain: it is again conserved in all three species, and the six surrounding residues are identical in yeast GAL7 and only one of them is missing in *E. coli* galT (Figure 2). In contrast, there is only spotty homology surrounding leucine-135, and the

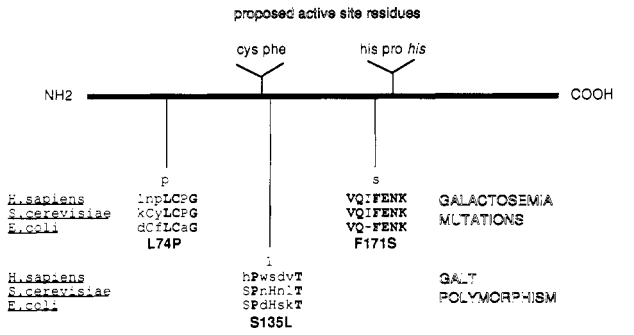

FIGURE 2: Homology analysis for two galactosemia mutations and one GALT polymorphism. The sequences of human GALT (Flach et al., 1990), yeast GAL7 (Tajima et al., 1985), and *E. coli* galT (Lemaire & Mueller-Hill, 1986) were aligned surrounding each substitution. Upper case residues are conserved in two species while boldface letters denote residues identical in all three organisms. The dash indicates a gap in the sequence alignments. The overall amino acid sequence identity for these three proteins is only about 35% (Flach et al., 1990). Putative active-site residues are indicated (Reichardt & Berg, 1988b), and the active-site nucleophile is italicized (Kim et al., 1990; Reichardt et al., 1992).

polymorphic residue itself is not conserved (Figure 2).

Discussion

We report here three amino acid substitutions in human galactose-1-phosphate uridyltransferase found in cell lines derived from two classic galactosemia patients. We sought to categorize these substitutions by expression analysis: two are disease-causing galactosemia mutations since the L74P and F171S mutations result in dramatic losses of GALT activity (Table I). In contrast, the S135L substitution is a polymorphism because it encodes normal GALT specific activity. The three GALT mutations reported here are all of the missense type, supporting the hypothesis of a preponderance of missense mutations in galactosemia (Banroques et al., 1983; Reichardt, 1989, 1991). All three substitutions are rare (Table II). Thus, these data confirm our hypothesis that galactosemia is heterogeneous at the molecular level (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992). Future studies on genotype–phenotype correlations will address a possible relationship between galactosemia mutations and biochemical phenotype. This kind of analysis proved very rewarding in the case of phenylketonuria (PKU; Okano et al., 1991).

The CpG dinucleotide is a hot spot for mutations in humans because they can undergo oxidative deamination of 5-methylcytosines (Cooper & Youssouffian, 1988). Two CpG mutations, R148W (Reichardt et al., 1992) and R333W (Reichardt et al., 1991), have been reported in galactosemia. Here we report a third substitution that could have arisen by oxidative deamination of a cytosine in a CpG dinucleotide: the S135L polymorphism (cf. Figure 1B). Thus, 3 of 11 reported GALT mutations, galactosemia mutations, and polymorphisms occurred at CpG dinucleotides. These three GALT mutations follow the rules predicted for deamination of 5-methylcytosine since they are all C to T transitions (Cooper & Youssouffian, 1988). It is noteworthy that one of the two galactosemia mutations reported here, L74P, creates a new CpG dinucleotide.

The "Black" variant of galactosemia, commonly found in African–American patients, is characterized by a substantial residual capability to metabolize galactose (Segal, 1989). In this context, it is interesting that the F171S mutation which was found in a cell line derived from an Irish–American individual (JR154) and has detectable residual activity (Table

I) also occurs in a line obtained from an African–American patient (GM2796). It could be either an ancestral mutation predating divergence of the races or a recurring mutation or be the result of genetic admixture which accounts for about one-third of all genetic markers in African–Americans (Reed, 1969). We favor the latter hypothesis since the F171S mutation appears to be cosegregating with the S135L polymorphism and, therefore, could be linked. Further characterization of African–American patients will examine these issues.

The overall amino acid sequence identity between human GALT (Flach et al., 1990), yeast GAL7 (Tajima et al., 1985), and *E. coli* galT (Lemaire & Mueller-Hill, 1986) is only about 35%. We had previously found that most galactosemia mutations occur in evolutionarily conserved domains, suggesting that they are of functional importance (Reichardt & Woo, 1991; Reichardt et al., 1991, 1992). In contrast, polymorphisms were found in variable areas, suggesting that they were less critical (Reichardt & Woo, 1991). The two mutations and one polymorphism reported here confirm this hypothesis since both mutations, L74P and F171S, occur in conserved domains while the polymorphism, S135L, affects a variable area (Figure 2). Furthermore, the F171S galactosemia mutation is located in proximity of the putative active-site nucleophile human histidine-186 (Reichardt et al., 1992). This is the third mutation affecting a residue in proximity of the proposed active-site nucleophile. The other two are the previously described Q188R and L195P mutations (Reichardt et al., 1991, 1992). Thus, galactosemia mutations highlight functionally important domains while GALT polymorphisms occur in less critical areas. It had been proposed that conserved patches of amino acid sequence identity might be helpful in structure–function considerations (Reichardt & Berg, 1988b), and our experience with the molecular analysis of galactosemia seems to support this proposition. Therefore, the molecular characterization of galactosemia mutations and GALT polymorphisms will provide valuable insights for a future rigorous biochemical structure–function analysis of human GALT. This study will involve an enzymologic examination of the reaction mechanism of the human enzyme in relation to functional domains of the GALT polypeptide. Mechanistic studies of the *E. coli* galT enzyme (Frey et al., 1982; Field et al., 1989; Kim et al., 1990) and delineation of functional domains in human GALT through an exhaustive analysis of patients will provide the necessary framework for this project.

In summary, we report here the molecular characterization of two novel galactosemia mutations, L74P and F171S, and one GALT polymorphism, S135L. Our data support the proposition that disease-causing mutations highlight functionally important residues while polymorphisms mark less critical amino acids. A similar pattern has been described in hemophilia B caused by factor IX deficiency (Bottema et al., 1991). Thus, we propose that the mutational analysis of galactosemia will provide some of the information necessary to begin a biochemical dissection of human GALT.

ACKNOWLEDGMENTS

We thank Stella Madu (Baylor) for EBV transformations. J.K.V.R. is an Associate and S.L.C.W. is an Investigator of the Howard Hughes Medical Institute.

Registry No. Gal, 59-23-4; GALT, 9016-11-9.

REFERENCES

Banroques, J., Schapira, F., Gregori, C., & Dreyfus, J. C. (1983) *Ann. Hum. Genet. 47*, 177–185.

Bottema, C. D. K., Ketterling, R. P., Ii, S., Yoon, H. S., Philips, J. A., & Sommer, S. S. (1991) *Am. J. Hum. Genet. 49*, 820–838.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Cooper, D. N., & Youssouffian, H. (1990) *Hum. Genet. 78*, 151–155.

Dobie, J. A., Holton, J. B., & Clamp, J. R. (1990) *Ann. Clin. Biochem. 27*, 274–275.

Field, T. L., Reznikoff, W. S., & Frey, P. A. (1989) *Biochemistry 28*, 2094–2099.

Flach, J. E., Reichardt, J. K. V., & Elsas, L. J. (1990) *Mol. Biol. Med. 7*, 365–369.

Frey, P. A., Wong, L. J., Sheu, K. F., & Yang, S. L. (1982) *Methods Enzymol. 87*, 20–36.

Kaufman, F. R., Xu, Y. K., Ng, W. G., & Donnell, G. N. (1988) *J. Pediatr. 112*, 754–756.

Kim, J., Ruzicka, F., & Frey, P. (1991) *Biochemistry 30*, 10590–10593.

Lemaire, H. G., & Mueller-Hill, B. (1986) *Nucleic Acids Res. 14*, 7705–7711.

Levy, H. L., & Hammersen, G. (1978) *J. Pediatr. 92*, 871–877.

Okano, Y., Eisensmith, R. C., Guettler, F., Lichter-Konecki, U., Konecki, D., Trefz, F. K., Dasovich, M., Wang, T., Henriksen, K., Lou, H., & Woo, S. L. C. (1991) *N. Engl. J. Med. 324*, 1232–1238.

Petry, K., Greinix, H. T., Nudelman, E., Eisen, H., Hakomori, S., Levy, H. L., & Reichardt, J. K. V. (1991) *Biochem. Med. Metab. Biol. 46*, 93–104.

Reed, T. E. (1969) *Science 165*, 762–769.

Reichardt, J. K. V. (1989) Ph.D. Thesis, Stanford University, Stanford, CA.

Reichardt, J. K. V. (1991) *Nucleic Acids Res. 19*, 7049–7052.

Reichardt, J. K. V., & Berg, P. (1988a) *Mol. Biol. Med. 5*, 107–122.

Reichardt, J. K. V., & Berg, P. (1988b) *Nucleic Acids Res. 16*, 9017–9026.

Reichardt, J. K. V., & Woo, S. L. C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2633–2637.

Reichardt, J. K. V., Packman, S., & Woo, S. L. C. (1991) *Am. J. Hum. Genet. 49*, 860–867.

Reichardt, J. K. V., Belmont, J. W., Levy, H. L., & Woo, S. L. C. (1992) *Genomics 12*, 596–600.

Segal, S. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, C. R., Beaudet, A. R., Sly, W. S., & Valle, D., Eds.) 6th ed., pp 453–480, McGraw-Hill, New York.

Tajima, H., Nogi, Y., & Fukasawa, T. (1985) *Yeast 1*, 67–77.

Waggoner, D. D., Buist, N. R. M., & Donnell, G. N. (1990) *J. Inherited Metab. Dis. 13*, 802–818.

Wood, W. I., Gitschier, J., Lasky, L. A., & Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1585–1588.